MOLECULAR STRUCTURE OF DEOXYRIBONUCLEIC ACID (DNA)

By R. LANGRIDGE,* W. E. SEEDS, Ph.D.,* H. R. WILSON, Ph.D.,* C. W. HOOPER,‡ M. H. F. WILKINS, Ph.D.,‡ and L. D. HAMILTON, D.M.§

(From the Wheatstone Physics Laboratory and the Medical Research Council, Biophysics Research Unit, King's College, London, and the Sloan-Kettering Institute, New York)

(Received for publication, June 19, 1957)

Although DNA is associated with complex biological phenomena and has many physicochemical properties not easily explained, there is no doubt that the molecule consists to a large extent of the regular, and in some ways simple structure (1–3), derived from x-ray diffraction. We will describe here briefly new data on the structure—the x-ray diffraction measurements and molecular model building being much more exact than in our previous publications (4, 5).

The DNA molecule has internal flexibility which enables the helical structure in the solid state to be twisted into either of two well defined configurations or changed reversibly from one to the other, *e.g.* (6). Proof of the correctness of the structure is doubly convincing because DNA can be studied in these two different forms. This proof resides in the fact that, with a base pairing of the Watson and Crick type (7), the helical structure may be arranged in these two ways, each of which is in agreement with the corresponding diffraction data.

We have already published some data on the *B* configuration (4) and discussed the nature of the *A* configuration (5). Here we give a preliminary account of more accurate data for both the *B* and *A* configuration. The procedure in this work is to build molecular models making use of interatomic bond lengths and angles obtained from measurements on relatively simple compounds and to calculate the x-ray diffraction intensities expected from such a structure. The model is then adjusted until the calculated diffraction corresponds with the measured x-ray intensities.

The B Configuration of the DNA Molecule:

When bound to protein in cells, DNA molecules exist in the *B* configuration (4). More information about the structure is now available: (*a*) because more exact measurements of x-ray intensities have been made, and (*b*) because new

* From the Wheatstone Physics Laboratory, King's College, London.

‡ From the Medical Research Council, Biophysics Research Unit, King's College, London.

§ From the Sloan-Kettering Institute, New York.

stereochemical data have enabled the molecular models to be built closer to the form likely to be correct. The previous x-ray measurements (4) using semi-crystalline Na DNA have been redone and their accuracy increased. Parts of the x-ray picture are diffuse, and the accuracy of the measurement in these regions is not as great as one would wish.

Crystalline DNA gives an x-ray picture composed of sharp spots, (Figs. 1 and 2) the intensity of which may be measured accurately. When the Na salt of DNA crystallises, the molecule does not retain the *B* configuration but changes into the *A* (6). However, we have found that the Li salt remains in the *B* configuration on crystallization and, using this material, we have made an independent and more accurate determination of the diffraction pattern of DNA in the *B* configuration. Comparison of the data shows only small differences between the configuration of the Li and Na DNA molecules, *e.g.* the Li molecule in the crystal is compressed along its length 3 per cent relative to the Na molecule. Another advantage of using the Li salt is that the Li ion scatters x-rays so much less than does the Na ion that it may be ignored when calculating the diffraction of the structure. The position of the ions in the structure is somewhat uncertain, and wrong placing of an ion as heavy as Na could introduce error into the calculation. The new stereochemical data used in building the model consists mainly of bond lengths and angles in a phosphate di-ester (8) and in thymidine (9).

Li DNA.—The microcrystals in a fibre of Li DNA are orthorhombic and, at about 66 per cent relative humidity, the dimensions of the unit cell are: *a* 22.72 ± 0.09 A, *b* 31.28 ± 0.11 A, *c* 33.60 ± 0.11 A. The axes of the helical molecules are in the *c* direction which is parallel to the fibre axis, and two molecules pass through the unit cell—one molecule being displaced relative to the other along this axis by about one-third of the helix pitch length. The distance between centres of molecules which make contact on the same level is 22.72 A, but where the contact is between molecules at different levels, the distance is as small as 19.33 A. Displacement of the molecules by one-third of the pitch length allows them (on account of being double and not single helices) to come more closely together than if they were on the same level. There are in the molecule exactly ten nucleotide pairs per turn of the helix. As the symmetry of the lattice reflects the symmetry of the molecule, one expects that if the ten-fold helices arrange themselves in a perfect crystalline manner, they will form an orthorhombic crystal.

When humidity is increased, the lattice swells and becomes less perfect, and the distance between centres of the molecules are more nearly equal than at 66 per cent humidity; at 75 per cent relative humidity the distances between centres of molecules are 22.8 and 24.4 A. At high humidity the helices arrange themselves in the same way as the Na, K, and Rb salts, *i.e.*, in a hexagonal form, with the molecules all equally separated by ~27 A between centres at

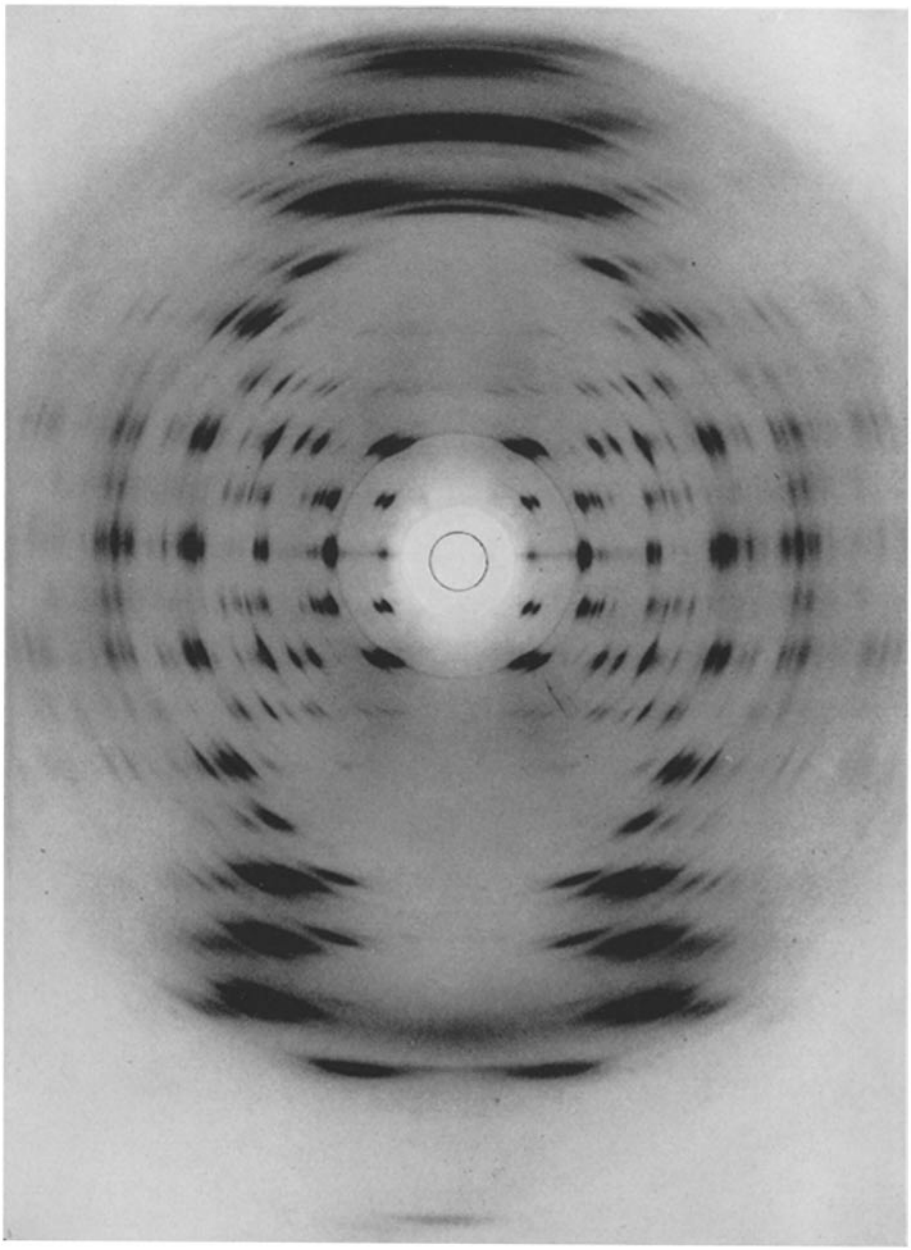

FIG. 1. X-ray diffraction photograph of crystalline Na DNA with the molecules in the *A* configuration at 75 per cent relative humidity.

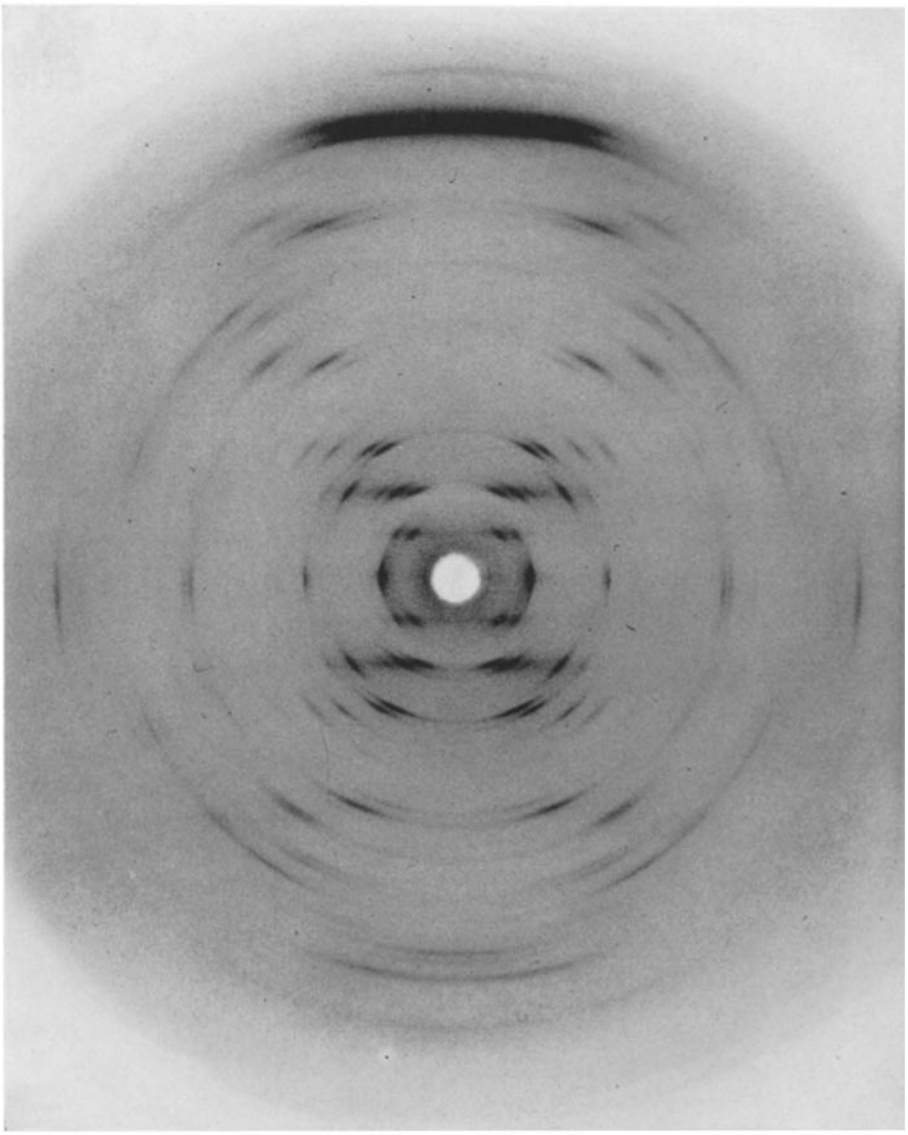

FIG. 2. X-ray diffraction photograph of crystalline Li DNA with the molecules in the *B* configuration at 66 per cent relative humidity.

92 per cent relative humidity, but the symmetry of the tenfold helix, however, does not permit the crystallization to be perfect.

The B Configuration.—In a general way the structure (See Table I, and Fig. 3) resembles that of Crick and Watson (7), and is a refined form of the structure

previously described by us (4). To fit the x-ray data, the atoms in the helix need to be placed on the whole as close as possible to the helix axis, and to help achieve this we have placed the base-pairs so that the helix axis passes centrally through each pair. The configuration of the base-pairs is not much different from that described by Crick and Watson (7). The planes of the bases are 3.36 A apart as is observed in crystalline Li DNA. The sugar ring is

TABLE I

DNA B-Structure Coordinates. Helix: 10-Fold Screw 3.36 A

36.0° Coordinates of Diad Axis (r, 0, 0)

We have taken the position of atoms in the nitrogen base rings as being approximately the same for the two purines and for the two pyrimidines.

Group	Atom	r(A)	ϕ (degree)	Z(A)	Group	Atom	r(A)	ϕ (degree)	Z(A)
Adenine	N_9	4.60	80.0	0		C_6	3.10	114.0	0
	C_8	4.80	95.0	0		O_6	3.00	139.0	0
	N_7	4.00	107.0	0		N_1	2.30	90.0	0
	C_5	2.70	97.0	0		C_2	3.30	72.0	0
	C_6	1.40	114.0	0		O_2	3.50	50.0	0
	N_1	0.90	42.0	0	Sugar				
	C_2	2.10	31.0	0		C_1	5.78	70.6	0.00
	N_3	3.20	48.0	0		O_s	6.37	70.6	1.32
	C_4	3.30	73.0	0		C_4	7.77	71.1	1.21
Guanine	O_6	1.80	158.0	0		C_5	8.39	78.8	2.18
	$(NH_2)_2$	2.90	5.0	0		C_3	8.12	73.5	−0.26
						C_2	6.84	74.0	−1.00
Thymine	N_3	4.60	80.0	0	Phosphate				
	C_4	5.06	95.2	0		P	9.41	58.7	−0.96
	C_5	4.50	111.0	0					
	$(CH_3)_5$	5.52	123.3	0		O_1	9.22	68.2	−0.88
						O_2	10.36	57.7	−2.07
						O_3	9.90	56.5	0.38
						O_4	8.13	53.1	−1.28

almost flat. The phosphate group has been built to conform with the data on dibenzyl phosphoric acid (8).

The structure is reasonable on stereochemical grounds, though some of the distances between atoms in contact are rather small. The O-C distances of 3.0 A are not objectionable and can probably be increased by further adjustment. For simplicity, in our model the base-pairs have been made planar and exactly perpendicular to the helix axis. This results in short distances between the C_2 atom of the sugar and the base of the neighbouring nucleotide in the chain. The distance to C_4 on a pyrimidine is 3.35 A (and 3.40 to the corresponding C_8 of a purine). To the centre of the methyl group on thymine the distance

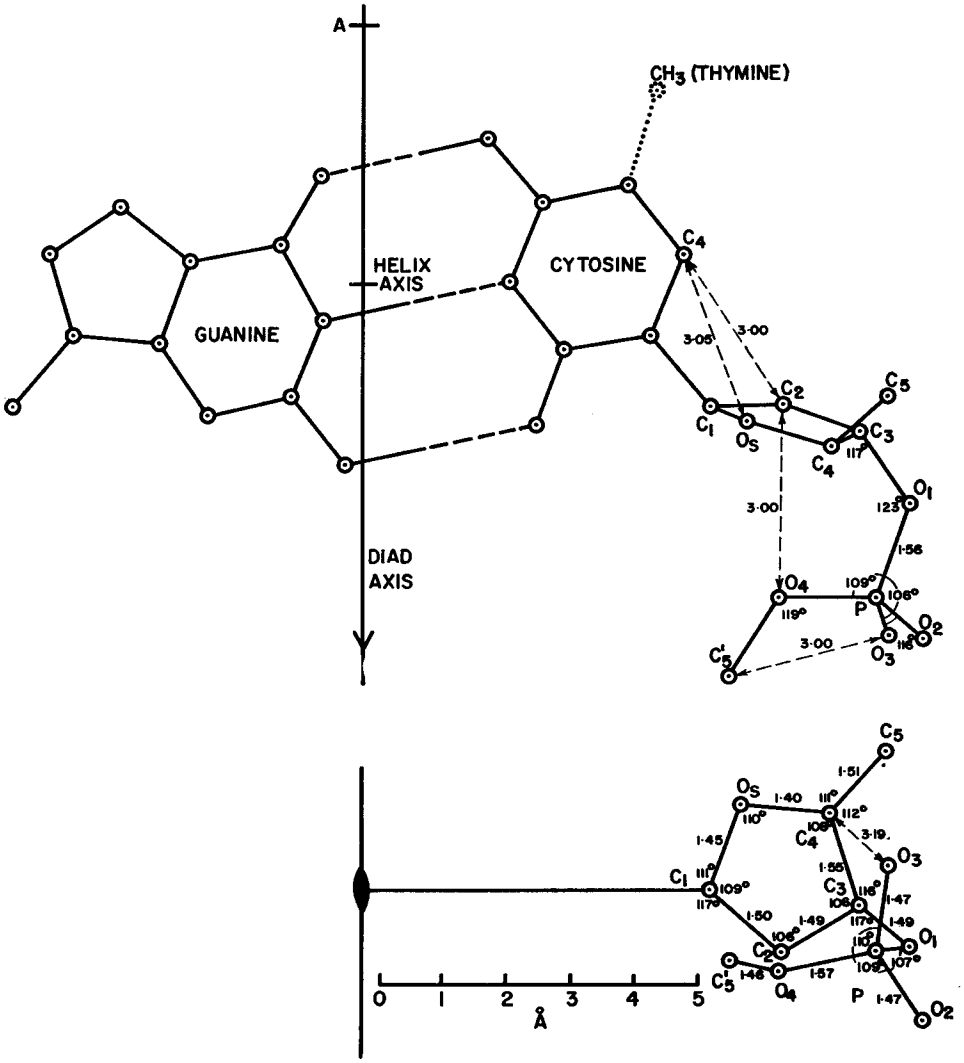

FIG. 3. The *B* configuration of DNA showing a base-pair and projections of a nucleotide on planes perpendicular to the helix axis and to the diad axis. Bond angles and lengths (approximate only) and short interatomic distances are shown. The position of the helix axis in the Watson-Crick model in relation to the base-pair is at *A*.

is about 3.1 A. (The position of this group is uncertain, and with Dr. M. Spencer we have re-estimated its position following the procedure of Pauling and Corey (10)). However, if the bases are twisted slightly about an axis at right angles to the helix axis and the diad axis, these distances increase considerably and, for a 5° twist, become about 3.5 A. This twisting would require little energy,

for the hydrogen bonds would be bent through less than 5°. The twist has little effect on the diffraction pattern and might be less than 5° for bases other than thymine.

Agreement of the B Model with the Diffraction Observed.—Figure 4 shows the observed x-ray intensity and the calculated diffraction of the structure postulated. The Fourier transform of the structure was obtained as in our previous work (4) except that the calculation was greatly accelerated by the use (in collaboration with Dr. M. P. Barnett of International Business Machines, United Kingdom Ltd.) of an I.B.M. 650 computer (11).

The x-ray intensities from the Li DNA have been corrected for the effect of the relative displacement of the DNA molecules along the helix axis. Some of the corrected intensities change noticeably if the value chosen for the displacement is altered slightly; other intensities are obtained by measurement of faint spots on the x-ray film or of spots overlapped by others. For these reasons many corrected intensities are not very reliable and are marked so in Fig. 4. Some of the main reliable corrected intensities obtained from the Na DNA are also plotted on the same scale and there is good agreement between the two sets of data. If the proposed structure is correct, the observed points will fall on the calculated curve. On the 5th and 6th layer lines, and on the zero and 3rd layer lines for values of ξ greater than 0.1 A^{-1}, the smooth curve represents the sum of the squares of two amplitude terms, and indicates in a very general way how the observed intensity varies. But, on account of interference between the two terms, the value of the square of the Fourier transform will fluctuate markedly as ξ varies, and hence, it has been calculated separately for each corresponding spot on the x-ray photograph.

In our calculation we have, for simplicity, assumed that the water molecules in the structure are placed between the DNA molecules, but not in exact positions. There is a moderate measure of agreement between the observed and calculated intensities in the outer part of the diffraction pattern and this shows that our assumption is probably not much in error. The agreement of intensities is, however, better in the central part of the pattern where the intensities are much less affected by the precise arrangement of the atoms.

The A Configuration of the DNA Molecule:

When the relative humidity of the atmosphere around a fibre consisting of the Na, K, or Rb salt of DNA is reduced from 90–85 per cent to 75 per cent, the DNA molecules change from the *B* into the *A* configuration. We have previously described in a rough way how in the molecule in this state the base-pairs are inclined at $\sim$70° to the helix axis and are displaced from it (5), and how the sugar-ring and phosphate group are at about the same distance from the helix axis. We also indicated very roughly how diffraction from such a molecule would resemble that observed.

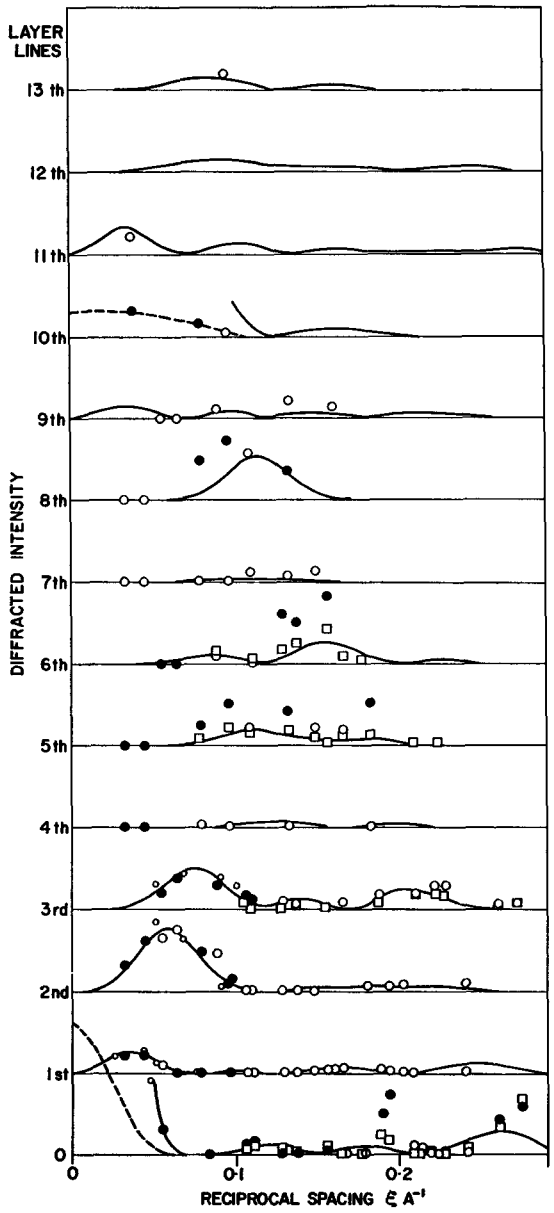

FIG. 4. Agreement between the observed and calculated intensities of diffraction for Li DNA in the *B* configuration. The more reliable observed values are shown by filled black circles and less reliable values are shown by unfilled circles of the same size. The small unfilled circles show values for Na DNA. The smooth curve is the sum of the squares of the calculated amplitudes, and where these terms interfere noticeably, the intensities have been calculated individually for each x-ray spot and are shown by square symbols. Intensities on the 10th layer line and the dotted curve on the zero layer line are plotted on 1/10 scale.

The coordinates for an *A* type structure are given in Table II. The same base-pairs are used as in the *B* structure described above, and the distances between atoms in contact in the phosphate-sugar chain are slightly greater than for the *B* structure, although 11 instead of 10 nucleotide pairs occur in one turn of the helix. Deviations from bond angles and lengths observed in simpler compounds are small; the average difference of angle is approximately 3°.

Contribution of the Metal Ions to the Diffraction of the Molecule.—We assume that in the crystal all space not occupied by nucleic acid is filled by water, and when allowance is made for x-ray scattering of the water, we find the Na ions

TABLE II

DNA A-Structure Coordinates. Helix: 11-Fold Screw 2.55 A

32.7° Coordinates of Diad Axis (r, 0, 0)

Group	Atom	r(A)	ϕ (degree)	Z(A)
Sugar	C_1	8.60	37.5	−1.90
	C_2	8.78	35.0	−3.27
	C_3	9.00	43.2	−4.06
	C_4	9.80	47.9	−3.00
	C_5	10.05	56.7	−3.18
	O_s	9.19	45.7	−1.64
Phosphate	O_1	9.82	41.7	−5.25
	O_2	10.18	37.5	−7.56
	O_3	7.92	44.3	−6.90
	O_4	8.83	28.8	−6.10
	P	9.05	38.5	−6.52
Base	N_1	7.18	36.2	−1.55

The planes of the bases are inclined at 70° to the helix axis.

in the salt of DNA effectively contribute to the scattering of x-rays at zero angle, an amplitude 1/9 that due to the DNA alone. We expect, therefore, that the scatter from the ions might appreciably modify the diffraction pattern of the DNA. The contribution of the ions may be found, without knowing their positions in the structure, by measuring the changes in the diffraction pattern when Na ions are replaced by Rb ions which scatter x-rays more strongly. We have achieved this by using the Na and Rb salts of DNA, which form crystals with the same lattice, and by assuming that the ions occupy, as one would expect, similar positions in the crystal. It appears that the ions are not in well defined positions in the crystal, and that their major contribution to the diffraction is in its central region. In calculations on the *A* and *B* configuration, when correcting for immersion in water, the effective number of electrons scattering x-rays per atom has been taken as follows: for base and sugar C =

3.2, N = 3.8, O = 4.3; for phosphate P = 9.6, O = 5.4; and for ions Na = 9, Rb = 32.

Agreement of the A Model with the Diffraction Observed.—Fig. 5 shows that there is reasonable agreement between the diffraction calculated from the

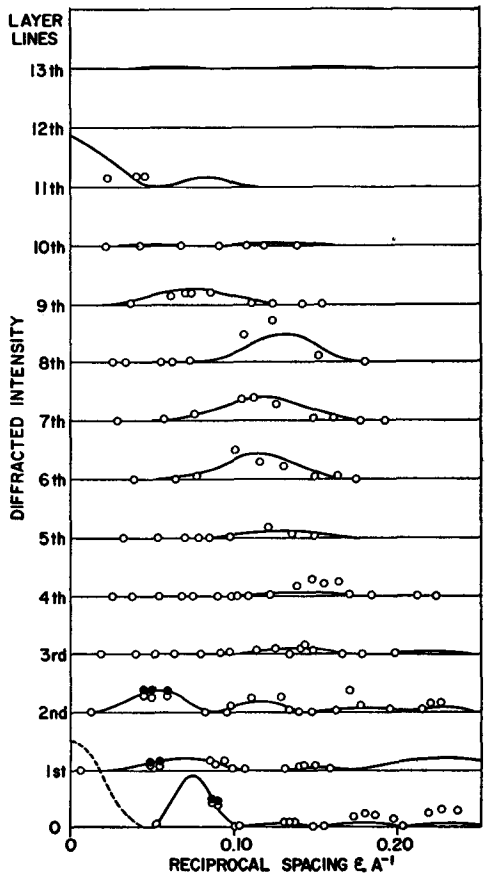

FIG. 5. Calculated and observed intensities of diffraction for DNA in the *A* configuration The observed points should lie on the calculated smooth curve. Agreement is improved when the values observed (unfilled circles) for Na DNA are corrected (filled circles) for the effect of the sodium ions. The dotted curve is plotted on 1/10 scale.

model (with no metal ions) and that observed. (The sum of the squares of the amplitude terms has been plotted and interference between them ignored at this stage). Agreement is noticeably improved, however, when the effect of the Na ions on the observed intensity is allowed for. The main discrepancy in the data is on the zero layer line for $\xi > 0.15$ A^{-1} where the calculated diffraction is too weak. This may be corrected by slight adjustment of the relative positions

of the phosphate groups in the polynucleotide chains and improved models are being built.

DNA Structure in Diverse Cells:

DNA's isolated from diverse cells engaged in a wide range of biological activities give the same x-ray diffraction patterns. Thus DNA in the resting state —the stored genetic material, isolated from sperm and bacteriophage, gives a pattern indistinguishable from that given by DNA isolated from very rapidly dividing cells, *e.g.* acute leukemic leukocytes, developing granulocytes, mouse sarcoma 180, and bacteria during the logarithmic phase of growth. These rapidly dividing cells have a high rate of both DNA turnover and protein synthesis. The same x-ray patterns are given by DNA's isolated from tissues that divide more slowly, *i.e.* with lower rates of DNA turnover, but in which there is considerable protein synthesis, *e.g.* DNA's prepared from liver, calf thymus, and lymphatic tissue. Finally, DNA from chicken erythrocytes—a DNA presumably metabolically inert—gives a similar picture. The identity of the x-ray pattern given by DNA's isolated from cells with such a wide range of biological activity suggests that DNA maintains the same double-helical configuration in cells whatever their activity.

CONCLUSION

We think it desirable to keep in mind the limitation of the x-ray method as applied to DNA (12, 13). Although x-ray diffraction tells that the helical structure is highly regular, and that this structure must hold for lengths of the molecule at least as long as 20 turns of the helix, it does not tell whether this holds for greater lengths, and gives little information about small chemical groups that might be irregularly attached to the helical skeleton. Moreover, the DNA studied is solid and as much as half of it may consist of non-crystalline material, about which the x-rays tell little. It is probably true to say, however, that other methods of studying DNA give little suggestion that DNA exists to any large extent in forms other than the double helix, or that irregularities, apart from base sequence, are a dominant feature of its structure.

We wish to thank Professor J. T. Randall, F. R. S., and Dr. C. P. Rhoads for encouragement, and Drs. A. R. Stokes and F. H. C. Crick for discussion. We are indebted to Dr. R. K. Barclay for preparations of DNA in particular from mouse sarcoma 180 and other mouse tissues and for K salts of DNA, and to Mr. D. A. Marvin for taking x-ray photographs which aided the determination of some of the diffraction intensities.

Two of us (W. E. S. and H. R. W.) wish to acknowledge the award of grants from the British Empire Cancer Campaign. The work at King's College was aided by a grant from The Rockefeller Foundation and that at Sloan-Kettering Research Institute by a grant from the Commonwealth Fund.

Computer time was helpfully provided by I.B.M. United Kingdom Ltd. and by I.B.M. World Trade Corporation.

BIBLIOGRAPHY

1. Watson, J. D., and Crick, F. H. C., *Nature*, 1953, **171,** 737.
2. Wilkins, M. H. F., Stokes, A. R., and Wilson, H. R., *Nature*, 1953, **171,** 738.
3. Franklin, R. E., and Gosling, R. G., *Nature*, 1953, **171,** 740.
4. Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K., and Hamilton, L. D., *Nature*, 1955, **175,** 834.
5. Wilkins, M. H. F., Seeds, W. E., Stokes, A. R., and Wilson, H. R., *Nature*, 1953, **172,** 759.
6. Franklin, R. E., and Gosling, R. G., *Acta Cryst.*, 1953, **6,** 673.
7. Crick, F. H. C., and Watson, J. D., *Proc. Roy. Soc. London, Series A*, 1954, **223,** 80.
8. Dunitz, J. D., and Rollett, J. S., *Acta Cryst.*, 1956, **9,** 437.
9. Huber, M., *Acta Cryst.*, 1957, **10,** 129.
10. Pauling, L., and Corey, R. B. *Arch. Biochem. and Biophysics*, 1956, **45,** 164.
11. Langridge, R., Barnett, M. P., and Mann, A. F., *Acta Cryst.*, to be published.
12. Wilkins, M. H. F., *Cold Spring Harbor Symp. Quant. Biol.*, 1957, **21,** 75.
13. Wilkins, M. H. F., *Biochem. Soc. Symp.*, 1957, **No. 14,** Cambridge University Press, 13.